DOCUMENT RESUME

ED 109 431 CE 004 276

AUTHOR Smedley, Rande H.; Olson, George H.
TITLE Graduate Follow-Up Studies: How Useful Are They?
PUB DATE 75
NOTE 28p.; Paper presented at the Annual meeting of the American Educational Researchers Association (Washington, D. C., April 1, 1975)

EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS *Comparative Analysis; Curriculum Planning; Decision Making; *Followup Studies; *Graduate Surveys; Longitudinal Studies; *Relevance (Information Retrieval); Research Methodology; *Statistical Bias

ABSTRACT

Follow-up surveys may fall prey to several sources of bias and error, among them lack of control over independent variables, lack of item validity and reliability, sampling biases, and observation bias. Two follow-up studies have been dissected to expose inherent limitations: the Texas Education Product Study (TEPS) and Project TALENT. The majority of the follow-up studies (i.e., studying graduates of a previous year) conducted fail to provide adequate information on which curriculum construction and planning can be based. Input is needed to aid in decision-making, however, and the best method readily available to provide this input is the follow-through approach (i.e., identifying a current class of students and following them beyond graduation). This approach comes closer to providing meaningful input by exercising, to some extent, a method of measuring relevant independent variables. Follow-up surveys, as they are commonly practiced, are not worth the time and effort spent conducting them. Follow-through surveys come much closer to producing information relevant to the relationships of concern to decision-makers and therefore prove much more worth the time and effort invested in using this approach. (Author/PR)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED109431

Graduate Follow-Up Studies: How Useful Are They?

Rande H. Smedley George H. Olson
Dallas Independent School District

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: CE EA
In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

Abstract

Some of the more salient limitations of survey research, in general, and of mail surveys, in particular, are discussed. Also discussed are two general approaches to conducting follow-up studies. The first of these (studying graduates of a previous year) is criticized as being inferior to the second (identifying a current class of students and following them beyond graduation). Suggestions are made for improving the validity, and thus the interpretability, of follow-up data.

U.S. DEPARTMENT OF HEALTH EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

CE 004276

Graduate Follow-Up Studies: How Useful Are They?[1]

Rande H. Smedley George H. Olson
Dallas Independent School District

Many reasons have been given to justify conducting follow-up studies of high school graduates. Usually it is purported that such studies enable school districts to evaluate the effectiveness of its instructional programs and curricula (Nelson [1964], Rothney [1964], and Kremer [1970]). The data collected in these studies are often used as input to administrative decision-making concerned with future planning and allocation of educational resources.

Among the types of data collected in graduate follow-up studies, those mentioned most frequently in the literature are the following (See Flanagan & Cooley [1966], Grant [1968], Snyder [1969], Kremer [1970] and TEA [1971]):

1. Comprehensive information on employment and educational status.
2. Information concerning social and geographical mobility.
3. Information concerning the extent to which pregraduate education and experience prepared graduates for their roles in business, industry, and society.
4. Information concerning the relationship of various high school aptitude and attitude factors to post-graduate success and satisfaction.
5. Information regarding graduates' perceptions of their formal educational programs and services.
6. Information bearing upon the process by which young people choose their life work.

The quality of information gathered in graduate follow-up studies varies considerably depending upon the type and quantity of information sought, the data collection strategy employed, the sampling plan, and a host of other factors related to the design and implementation of the

---

[1]Paper presented at the annual meeting of the American Educational Researchers Association, Session 13.09. Washington D.C., April 1, 1975.

study. Obviously, studies differ in terms of their methodological soundness, and this has important implications for the validity of the data. Even in those studies that are better designed, however, it is important to ask the question: Do the data collected in graduate follow-up studies ever provide the information necessary to give unbiased answers to the questions asked? We think the best answer to this question is generally "no", but that it is possible to design studies which provide more useable information than others. In this paper we will describe some of the more serious limitations of follow-up studies and research surveys in general. Two approaches to follow-up studies will be scrutinized, including examples, and what we consider the most valuable approach will be discussed.

Background Considerations

Before describing follow-up studies in detail, it is perhaps profitable to provide some background information on that species of research of which follow-up studies are examples. In this section, some of the more saliant limitations of survey research in general, and of mail surveys in particular, are discussed. In a later section we will discuss two general approaches to conducting follow-up studies. The first of these approaches (studying graduates of a previous year) is criticized as being inferior to the second approach (identifying a current class of seniors and following them through successive years following graduation).

Ex Post Facto Research

Most follow-up studies can be methodologically classified as *ex post facto* research, research that is completed after the fact. This type of

research may be characterized by a lack of formal control over the independent variables of interest. A quote from Kerlinger (1964) provides a good definition.

> "That research in which the independent variable or variables have already occurred and in which the researcher starts with the observation of a dependent variable or variables. He then studies the independent variables in retrospect for their possible relations to, and effects on, the dependent variable or variables [p. 360]."

The major danger confronting *ex post facto* studies is generally known as the *post hoc fallacy*, after this, therefore because of this. This fallacy refers to the tendency to assert that one thing causes another simply because it precedes the other in time. For instance, it may be all too easy to conclude that one high school has a better curriculum than another high school simply on the basis of an observation that the graduates of the first high school have a higher median income than graduates of the second high school. However, without adequate data it would be impossible to rule out such competing hypotheses as, "The students in the two schools come from families with vastly different educational and socioeconomic backgrounds." If the data were known, it might be concluded that the curriculum in the second high school was actually superior to that in the first school. This might be the case, for example, if it were known that the expected median income of students with backgrounds similar to those in the second school was appreciably lower than that observed, while median income of graduates with backgrounds similar to those from the first school was about the same as that observed. In other words, the difference in median income may or may not be due to differences in curriculum between the two schools. To answer this question, additional data would be needed. At the very minimum, some method of controlling for differences in student

background (and a host of other variables as well) would be needed before generalizations regarding this effect of curriculum could be made.

The lack of control over independent variables and, usually, lack of randomization are major deficiencies inherent in all *ex post facto* research. Although it is often possible, in theory, to draw an initial random sample, it is not possible either to assign subjects to groups at random or to assign treatments to groups at random. Rather, the researcher must be content to study groups of individuals who have been naturally assigned because they possess differential characteristics and traits which are usually only extraneously related to the objectives of the study. The danger in this is that these "variables of self-selection" often influence or are otherwise related to the variables being studied.

Consider an over simplified example, suppose that a researcher is interested in the effect of attitude toward school on on-the-job success. To study this, he sends out a questionnaire to a sample of graduates five years after graduation. The questionnaire contains the following two items.

"What is your present income?"

"Did you enjoy going to high school?"

Suppose he sends out 500 questionnaires of which 335, or two-thirds, are returned. After deciding upon some criterion of success based upon income, he constructs the following fourfold table.

	Successful	Unsuccessful	Total
Enjoyed School	47% (156)	20% (68)	67% (224)
Did Not Enjoy School	23% (78)	10% (33)	33% (111)
Total	70% (234)	30% (101)	

The first, albeit naive, interpretation might be (1) that there were more than twice as many successful graduates as unsuccessful graduates, (2) that more high school graduates enjoyed school than did not enjoy school, and (3) that success is largely dependent upon enjoying school. An obvious implication would be to make school as enjoyable as possible for students!

Now, consider some reasonable alternative explanations which could account for these results. It is well known that the probabilities associated with the return of the questionnaire are not equal for all members of a sample selected at random for a given population (Donald [1960]). On the contrary, there is evidence to suggest that the chances of a returned questionnaire are best if the respondent is a successful individual of better than average intelligence (Shuttleworth [1941], Reuss [1943], and Rothney & Mooren [1952]). In short, it is reasonable to expect a higher rate of return from more intelligent, successful people. The data in the table above merely reflect this line of reasoning. In constructing the data for the table, we assumed (not unreasonably) that the bulk (75%) of the non-respondents were "unsuccessful" individuals. Without knowing this beforehand, however, any

conclusions based on these data would likely be very misleading.

Suppose it were possible to go back and, through some, usually expensive, means, obtain data from the initial non-respondents. The fourfold table might now look like the following.

	Successful	Unsuccessful	Total
Enjoyed School	37% (184)	30% (150)	67% (334)
Did Not Enjoy School	18% (91)	15% (75)	33% (166)
Total	55% (275)	45% (225)	

What conclusions can now be drawn? For one thing, it cannot be claimed that there were twice as many successful graduates as unsuccessful graduates as could have been implied prior to collecting the additional data. As before, however, it can be concluded that two-thirds of the graduates perceived school as having been enjoyable. Also, there is still a relationship, though considerably attenuated, between success and perceived enjoyment of school. Can it now be concluded that an enjoyable school experience is likely to lead to success later in life? The answer must still be no. The reason is that only the respondents' current perceptions of an experience that occurred five years earlier have been measured. The course of time can have many effects on memory traces of experiences. Many additional assumptions would have to be made before the present perceptions of the respondents can be accepted as true reflections of their perceptions while in high school. Many people, for good reasons, may not be willing to make such assumptions.

An alternative hypothesis, also supported by the data given above, is that whether or not school is perceived as having been enjoyable depends upon the amount of success the individual attains later in life. The best approach, however, would be to forego any causal interpretations and simply conclude that successful individuals are more likely to perceive their school experiences as having been enjoyable than are nonsuccessful individuals. Clearly, this does not lead to any better understanding of the effect of attitude toward school on success than was available prior to beginning the investigation. With foresight, a major improvement in the study could have been made by collecting attitude-toward-school measurements on the sample while they were still in school. It should be noted, however, that even though this might have improved the study greatly, any conclusions would still have to be modified to take into account the possibility that attitude toward school, while the student is still in school, is probably influenced by several other determinants of success (e.g., intelligence, parents' educational level, etc.). Thus, although the study would be improved by collecting earlier data, it would still be an *ex post facto* study and thus susceptible to the *post hoc fallacy*.

If the collection of one piece of earlier data can improve the study, then it might be expected that the collection of several pieces of earlier data would improve the study even more. Indeed, if the right data are collected, it would open the possibility of statistical control of irrelevant variables. Thus, powerful statistical methods could be brought into play in an effort to ferret out possible causal relationships.

Survey Research

Kerlinger (1964) has distinguished survey research from *status surveys*. Status surveys tend to be routine fact and data gathering activities. In education, status surveys are usually conducted as a means of gathering facts about the schools such as teacher-pupil ratios, curriculum practices, median income of graduates, availability of facilities, and so forth. The purpose of these surveys is generally to determine present status on variables of interest and not to study interrelations among these variables.

Survey *research*, on the other hand, is defined as a branch of social-scientific investigation. As such, it carries with it a connotation of scientific rigor. In contrast to the status survey, survey research is usually focused upon uncovering relationships among important variables. For example, a survey researcher might seek an answer to the question, "Is attitude toward school related to post-graduation success?" As was seen in the previous section, the answer to this question requires the utilization of considerably more sophisticated techniques then merely measuring current success and attitude-toward-school status of graduates.

Status surveys are usually considerably easier to design than are studies in survey research. Generally they have the appearance of being routine fact-gathering exercises. This may often lead to a feeling that the data collected in a status survey are relatively free of bias and other sources of error. Unfortunately, the opposite is more often the case. Echternacht (1973) discussed some implications resulting from reliability and validity studies conducted with an

educational survey. A "resurvey" technique was used in which a subsample of participants in the National Longitudinal Study* (NLS) were sent a second survey for the reliability study. Parents of a second subsample of NLS participants received a survey questionnaire for the validity study. Echternacht (1974) summarized his results as follows:

> In the resurvey conducted for a subsample of NLS participants, significant biases were found in items relating to father's occupation and parent's income--two items comprising the derived, composite SES scale. Two four-item social-psychological scales were found to have reliabilities that would likely reduce the multiple correlation between the two scales and a reading comprehension test score. In collecting and processing data, deficiencies in the data collection systems, such as the lack of follow-up on the educational plans and aspirations question, were uncovered. (p.11)

It seems that routine fact-gathering is not as free from problems as may be commonly thought.

Mail Surveys

Several types of survey research can be identified. These include the personal interview, telephone surveys, and mail surveys. Personal interviews are generally considered, by most experts, the best examples of survey research. As the name implies, this type of survey calls for

---

*For a detailed discussion of the NLS see Hilton, T. L., Rhett, H., Broudy, I. L., Bower, C., Carter, M., Creech, R., & Echternacht, G. J. The base-year survey of the national longitudinal study of the high school class of 1972. Princeton, N. J.: Educational Testing Service. Final report for contract No. OEC-0-72-0903, 1973.

a face-to-face encounter between the respondent and a trained interviewer. Unfortunately, the power of the personal interview as a data gathering technique does not come cheaply. Interview schedules are difficult to construct and interviews tend to be time-consuming and expensive. If the area in which the survey is to be conducted is large (e.g., a whole state, the Nation) the cost of conducting personal interviews will be greatly increased. Jackson & Rothney (1961) for instance, found that for every one dollar spent on the mail questionnaire technique, sixty dollars would have been spent using the personal interview technique.

Telephone surveys suffer from an inability to obtain detailed information. Since the interviewer is usually unknown to the respondent, he faces possible nonresponse, uncooperativeness, and a reluctance to answer anything other than simple, superficial questions (Kerlinger, 1964). Another serious difficulty with telephone surveys is the possible sampling bias that is likely to result from the fact that not everyone has a telephone and the fact that more women are likely to be reached than men.

The mail survey has the advantage of being able to collect large amounts of data from large segments of the population inexpensively and for this reason enjoys wide popularity in areas where large sums of money are not available for research. Beyond the advantage of being inexpensive, however, there is little to recommend its use, although Droege and Crambert (1965) express a contrary opinion. One deficiency inherent in mail surveys is the usual high rate of nonresponse. The difficulty in obtaining responses from significant numbers of the intended

sample has led at least one prominent researcher to conclude that, "these defects are serious enough to make the mail questionnaire worse than useless [Kerlinger, 1964; p. 397]." Another researcher, Parton (1950) has similarly concluded that, "most mail questionnaires bring so few returns, and these from such a highly selected population, that the findings of such surveys are almost inveriably open to question [p. 400]."

The basic problem with mail surveys is that they are more susceptible than other types of survey methods to the many sources of bias to which survey research is particularly vulnerable.

Sources of Bias
In Survey Research

Bias may be simply defined as the deviation of a *survey* value from the *true* value in the population. This should not be confused with measurement errors or sampling errors which are defined as deviations of a sample value from the value which would be obtained by surveying the total population under the same conditions used in the sample (i.e., survey value).

Leslie Kish, a nationally recognized authority on survey sampling techniques, has categorized the many sources of survey errors by differentiating first between biases and variable errors and then between sampling and nonsampling errors (Kish, 1965). The first distinction, that between variable errors and biases, partitions the total source of error, defined as the deviation of the observed sample value (X_o) from the true population value (X_t) into two types of deviations: the deviation of the observed value from the expected sample value (X_s),

and the deviation of the expected sample value from the true population value. This distinction can be stated in a simple equation as

$$(X_o - X_t) = (X_o - X_s) + (X_s - X_t).$$

The first term on the right side of the equation represents those sources of sampling and nonsampling errors which are usually assumed to occur randomly in samples from the population being studied. These include errors which stem from the crudeness of the survey instrument, defects in quality control of data processing, faulty sampling procedures, and the use of imprecise statistical estimators. Usually these sources of error are dealt with, more or less routinely, by improving the mechanical aspects of the survey design.

The last term in the equation above represents sources of error due to bias and are much more difficult to deal with. As with the errors identified by the first term on the right side of the equation, biases can also be differentiated into sampling and nonsampling sources. Sampling biases stem, for instance, from problems with the sample-selection procedures, such as when members of the population have unequal selection probabilities. Sampling biases can also arise as a result of using consistent but biased statistical estimators. It is often possible to correct sampling biases by weighting subsections of the sample by their respective selection probabilities and by designing methodologically "tight" sampling procedures.

Nonsampling biases, on the other hand, represent a profound problem for the survey researcher. This problem arises partly out of the fact

that nonsampling biases effect the population survey value as well as the sample value. Thus, no matter how methodologically sound the sampling plan may be, there is the ever present threat that the sample value may not closely resemble the true population value. Kish (1965) has stated that this difference, "delineates the gap between the defined survey objectives and the entire set of survey conditions [p. 520]." What Kish is pointing out with this statement is that it is often difficult (if not impossible) to obtain estimates of the true value of many variables in the population (e.g., attitudes, opinions, beliefs, etc.). The reason for this is that often the procedures used to estimate these values do so only indirectly. Furthermore, there is an effect, well known to scientists, and particularly applicable to survey research, that tends to preclude our ever being able to estimate the true value. This effect, known as the *Heisenberg Principle*, essentially states that in the act of measuring something, that which is being measured is altered in some way. In other words, to measure the status of a variable, the status of that variable must first be interrupted. If the purpose of the study is to measure a student's attitude toward school, for example, his status quo must be obtained either by asking him a question directly or by requesting him to complete a questionnaire. In doing so, either his attitude toward school or his attitude toward responding to the query (from an objective research point of view, this distinction is immaterial) may be altered. The student, for instance, may now object to being asked to make public his attitude, or he may attempt to respond in a way he thinks would make the researchers happy. In either event, it may be impossible to be sure of his true attitude toward school .

Nonsampling biases arise from two major sources: bias due to observation and bias due to nonobservation. Observation bias can only occur when the intended respondent actually responds, in some recordable way, to the survey procedures. Bias enters when these responses are either given or recorded in some way that consistently misrepresents the respondents' true scores on the variables of interest.

Bias resulting from nonobservation, on the other hand, occurs when individuals, who for some reason are not included in the *effective* survey sample, are in some relevant way different from those who are included in the sample. When this occurs, estimates of population values based upon the sample values can be very misleading. An example of this type of bias was presented earlier in this paper. To illustrate the effect of bias that might result from nonobservation more dramatically, assume that a researcher is interested in determining the average number of years of post-graduate education attained by high school graduates. He chooses to determine this by conducting a mail survey in which graduates of five years ago are asked to indicate the number of years of schooling they have attained since graduating. Now, let

p represent the proportion of individuals who return the questionnaire,

$\bar{Y}_R$ represent the mean number of years of post-graduate education among those who return the questionnaire, and

$\bar{Y}_N$ represent the mean number of years of post-graduate education among those who do not return the questionnaire.

The mean number of years of post-graduate education in the entire sample

can be represented by the equation

$$\bar{Y} = p\bar{Y}_R + (1 - p)\,\bar{Y}_N.$$

Without much difficulty, it is easy to suppose that those individuals who have had post-graduate education would be more likely to return a questionnaire dealing with post-graduate education than would those individuals who have not had this experience. Thus the following values might not be unreasonable.

$$p = 0.62$$

$$\bar{Y}_R = 2.9$$

$$\bar{Y}_N = 0.6$$

The observed response, $\bar{Y}_R$, taken without question, would lead to the conclusion that the graduates are indeed highly educated. However, if all the data were known,

$$\bar{Y} = (.62)\ (2.9) + (1 - .62)\ (0.6),$$

or $\bar{Y} = 2.02$, almost a full year's difference. Thus, the estimation based only upon the returned questionnaires would have been grossly exaggerated. The data collected by Shuttleworth (1941) provides a concrete example of the kind of bias that can result from non-observation. Using a mailed questionnaire, he received responses from 56 percent of his initial sample of graduates. A tabulation of responses indicated .5 percent unemployment among those responding to the mailed questionnaire. Through telephone calls and personal interviews, Shuttleworth was later able to collect information from an additional 38 percent of his initial sample of graduates. A tabulation of responses indicated 5.6 percent unemployment among respondents to phone calls and interviews. Had Shuttleworth terminated his study after collecting mailed questionnaire data alone, he would have been omitting a larger percentage of data relevant to unemployed graduates.

There are several ways in which nonobservation of various segments of the survey population can occur. In the mechanics of selecting the survey sample, some elements in the population may be unwittingly excluded. It can also result from faulty execution of the survey procedure. Nonobservation also occurs when, for one reason or another, individuals in the survey sample fail to return completed questionnaires. This latter problem (nonresponse) poses the greatest threat for mail surveys. Reasons for the nonreturn of completed questionnaires span the range from inability to locate individuals in the first place, to individuals' refusals to answer the questionnaire. That nonresponse is indeed a threat is evidenced by the preponderence of mail surveys reported in the literature in which the researchers failed to obtain data on half or more of their intended sample (Grant [1968], Snyder [1969], and Kremer [1970]). We have also received a low response rate (33 percent) in a graduate follow-up study we are conducting in Dallas. This low response rate is in spite of the fact that we have employed extensive data collection procedures including three waves of mailed questionnaires followed by reminder cards and several address validation techniques. It should be pointed out, however, that not all mail surveys are doomed to low response rates (Shuttleworth [1941], Rothney & Mooren [1952], and Donald [1960]). Several especially well-designed studies have elicited much higher response rates--sometimes as high as 80 and 90 percent (for examples, see Kish, 1965; p. 538). From her review of the literature, Kish arrived at the following conclusions in regard to mail surveys:

> high response rates are more likely to be obtained with brief questionnaires,
>
> several mailings will often incurese the response rate over 80 percent,
>
> interviews of a subsample of nonrespondents will further increase the response rate, and
>
> low responses to the first or even second mailing should not be accepted since they are likely to contain severe selection biases.

Eckland (1965) described several techniques for increasing return rates. These include (a) sending the questionnaire to nonrespondents periodically, (b) following the questionnaires with reminder cards or letters, (C) sending the questionnaires by certified mail with return receipts requested, and (d) telephoning nonrespondents to ask them to respond.

Types of Follow-Up Studies

Two Examples

In this section, we discuss two previous follow-up studies which we believe are representative of basically two different approaches for the conducting of follow-up studies. The first of these, the Texas Education Product Study (TEPS), was commissioned by the Texas Education Agency (TEA, 1971) to obtain data on individuals who left Texas high schools during the years 1963-1964 and 1968-1969. Data for this study were apparently collected during 1971 and 1972. The second study of interest, Project TALENT (Flanagan and Cooley, 1966), is actually a continuing follow-up study of a national sample of students who graduated or left high school during the years 1960-1963.

There are several comparisons which can be drawn between the TEPS and Project TALENT studies which have important bearing on considerations for conducting follow-up studies. The first comparison concerns the rates of return of questionnaires obtained in the two studies. Both studies included samples of prior students for whom available addresses were from one to five years old. On the assumption that the TEPS sampling plan included approximately equal numbers of students from each of the two populations of the study, it is possible to estimate their rate of return for the 1968-1969 sample to be about 48 percent and the rate of return for the 1963-1964 sample to be about 32 percent. Project TALENT, on the other hand, was able to elicit a 69 percent of return from the sample of students whose addresses were one year old and a 37 percent

rate of return from a sample of students whose addresses were five years old. It is likely that these differences in rates of return are due to the fact that the Project TALENT samples were actively involved in the purpose of the study while still in school. In the Project TALENT study, a considerable amount of testing was done in order to collect data that could later be used in ferreting out meaningful relationships in the data from the follow-up studies. Since both the TEPS and Project TALENT investigations utilized extensive data collection procedures, it seems reasonable to conclude that the initial pre-involvement of the Project TALENT students in the beginning stages of the study tipped the scale in favor of a higher rate of return from their sample.

A second important comparison between the TEPS and Project TALENT studies concerns the way nonrespondents were treated. It is useful here to define more clearly the *population* of nonrespondents as those individuals who are not included in the effective survey sample due to the ineffectiveness of the operational survey procedures. Thus, included in this population of nonrespondents are those individuals who are not found as well as those who fail to return the survey questionnaire even after receiving several reminders and other pleas to do so. For these latter individuals, the survey procedures simply fail to elicit any response.

When the population of nonrespondents is defined in this way, it is necessary, if data are to be collected from these individuals, to design a different set of data collection procedures. In the TEPS investigation, no further attempt was made to obtain data from nonrespondents. In the Project TALENT study, however, a special form of the questionnaire and a

different data collection procedure were used to obtain data from five percent of the nonrespondents. To insure that data obtained from this new sample would not contain biases due to nonobservation, project personnel invested in considerably more expensive techniques and were rewarded by a 93%-99% rate of return from the sample of nonrespondents. The advantage of the Project TALENT approach is obvious. It affords researchers the opportunity to measure the extent of bias present in the survey sample.

The final comparison is perhaps the most important. This comparison is concerned with the potential value of the two approaches, (i.e., the potential use to which the data collected in the follow-up survey may be put). In the TEPS project, it was not possible to obtain data on a variety of variables while the individuals in the sample were still in school. All data in that study were collected at one point in time. Because of this, the study results are limited to descriptive accounts of status. It is not possible, from the results presented in the TEPS report, to identify possible causes for these status indices.

Some examples from the TEPS report will serve to illustrate the point. In that study (TEA, 1971b; p. 4-1), it is reported that 65.2 percent and 47.9 percent of the respondents felt their high schools had done an effective job in preparing them for further education and work, respectively. This result is interesting, but it does not provide any information as to how the schools might better serve the 14.2 percent and 23.5 percent who felt that the high school was ineffective in preparing them for further formal education or work. There are, no doubt, many complex relationships among ability, aptitude, attitude, and cur-

riculum variables that contribute to students' post-graduation perceptions of high school effectiveness.

In another place, respondents were asked to respond to an open-ended question concerning school improvement. The most common criticism voiced by the respondents was, "Teachers should show more concern for students as individuals." This criticism, however, was obtained from only two percent of the respondents. How should this result be interpreted? Without knowing something more about these individuals while they were students, this question cannot be answered. It might be useful to know, for instance, whether these students tended to be "loners" or low-achievers, etc.

Finally, the TEPS study reported that the majority of respondents approved of extracurricular activities. Again, this finding is interesting, but what can be said in regard to the effect of participation in these activities on post-graduation success? Earlier data on such variables as level of participation and number of activities engaged in might have allowed recommendations, for example, as to which activities should be given more stress than others.

In contrast to the constraints that restricted personnel in the TEPS investigation, researchers for Project TALENT enjoyed the luxury of being able to design a follow-up study well before students in the designated population were to graduate. Thus, these researchers had the opportunity to research the educational literature and select those variables that held the greatest potential value for scientific research. Because the students in their sample had not yet graduated, Project TALENT researchers

were able to obtain valuable baseline measures on these variables. This, in turn, eventually allowed them and other researchers to derive important empirical relationships between processes that occur in school and post-graduation experience. Although it must still be classified as *ex post facto* research, the Project TALENT study has gone a long way toward understanding the effects of schooling.

Two Approaches to Conducting Follow-Up Studies

The follow-up studies conducted by TEPS and Project TALENT are characteristic of two types of approaches that might be taken to obtain information from high school graduates. The TEPS study was a traditional follow-up study in the sense that graduates from a previous year were studied. In contrast to the traditional follow-up approach, the Project TALENT study used what we will call a follow-through approach which identified a current class of seniors and followed them through successive years. The relative merits of these two approaches are discussed below.

Follow-up approach. The previous sections of this paper should be demonstrated that this approach severely limits the ability to identify causal relationships between within-school experiences and post-graduation expericences.

There are a number of specific problems which will be encountered in using this approach to study graduates. The follow-up approach requires that a specific class (or classes) be identified as the population of interest. In order to select a truly representative sample from this population, the relative propabilities of inclusion for the students in the sample would have to be known. Previous experiences in trying to locate names and (then current) addresses of students of several years ago suggest that it is not enough to assume

that these probabilities are equal.

Even if the complete roster of names and (then current) addresses can be located, there still remains the problem, in many cases, that the addresses are no longer current. Because a change in address is not likely to be independent of the objectives of the survey, the final effective sample will most liekly be severely biased. For example, it is reasonable to expect changes of address to be easier to obtain among the more educated individuals in the sample. Their parents/guardians are more stable and are likely to be able to supply current addresses. They are more likely to leave forwarding addresses when they do move. Thus, the less educated and less successful individuals may be more difficult to find due to high mobility and lack of forwarding addresses and, ultimately, may not be included in the survey's findings. Another group of individuals that may end up being "selected-out" of the study is that composed of married females. The resulting change of name may make it virtually impossible to find many women graduates. Since marriage is related to such variables as success on the job and years of higher education, failure to include these married women in the sample may seriously distort the survey results.

Perhaps the major disadvantage with the follow-up approach is that it does not provide decision-makers with the kind of information needed for such tasks as curriculum revision, resource allocation, and determining the importance of a wide variety of affective and cognitive variables to success and failure in contemporary society. This would require a host of earlier measurements on the individuals in the sample.

Follow-through. The comparisons drawn earlier, between

the TEPS and Project TALENT studies, serve well to point out the advantages of using a follow-through approach. This approach is particularly attractive to those Districts which have accumlated, to some degree, a longitudinal data base. Once the sample is choosen, it would be a fairly simple matter to collect additional relevant observations and measures from the students and their teachers. This information could later be used in drawing meaningful conclusions from the results of the follow-through surveys. Of course, this additional advantage would be obtained at an increased expense of developmental effort, but the potential usefulness of the results would surely make it worthwhile.

In addition to the increase interpretability of survey results, there are several additional advantages to be gained from the follow-through approach. Since the students will still be in school at the inception of the study, they can be conditioned to respond to the future follow-up questionnaires (see Krueck [undated] for a discussion of several procedures that might be used to encourage students to be future respondents). Furthermore, by sending out yearly address queries, it will be possible to maintain current address records. This technique, with slight modifications, might also be used to maintain respondents' interest in the survey. The address queries could be included along with a special-edition student newspaper from the respondents' high schools or a "fact-sheet" giving simple frequency data on student migration patterns.

Conclusions

It has been demonstrated that follow-up surveys may fall prey to several sources of bias and error, among these lack of control over independent variables (post hoc fallacy), lack of item validity and reliability, sampling biases (which include non-observation bias), and observation bias (Heisenberg Principle). Two follow-up studies have been dissected to expose inherent limitations. It is our opinion that the majority of the follow-up studies conducted fail to provide adequate information on which curriculum construction and planning can be based. Input is needed to aide in decision-making, however, and the best method readily available to provide this input is the follow-through approach of which Project TALENT is an example. This approach comes closer to providing meaningful input by exercising, to some extent, a method of measuring relevant independent variables. Follow-up surveys, as they are commonly practiced, are not worth the time and effort spent conducting them. Follow-through surveys, on the other hand, come much closer to producing information relevant to the relationships of concern to decision-makers and therefore prove much more worth the time and effort invested in using this approach.

References

Donald, M. N. Implications of nonresponse for the interpretation of mail questionnaire data. Public Opinion Quarterly, 1960, 24, 99-114.

Droege, Robert C. and Crambert, Albert C. Follow-up techniques in a large-scale test validation study. Journal of Applied Psychology, 1965, 49 (4), 253-256.

Echternacht, Gary J. Reliability and validity of questionnaire items---status, social-psychological and educational plans and aspirations. Educational Testing Service, Princeton, N. J. RB-74-12, April 1974.

Eckland, Brucek. Effects of prodding to increase mail-back returns. Journal of Applied Psychology, 1965, 49 (3), 165-169.

Flanagan, J. C. and Cooley, W. W. Project talent: one year follow-up studies. School of Education, University of Pittsburgh, 1966.

Grant, Claude W. A follow-up study of spring 1966 high school graduates in the state of utah. Personnel and Guidance Journal, 1968, 47 (2), 157-162.

Jackson, Robert M. and Rothney, J. W. M. A comparative study of the mailed questionnaire and the interview in follow-up studies. Personnel and Guidance Journal, 1961, 39 (7), 569-571.

Kerlinger, F. N. Foundations of Behavioral Research. New York: Holt, Rhinehart, and Winston, Inc., 1965.

Kish, Leslie. Survey Sampling. New York: John Wiley and Sons, 1965

Kremer, Bruce J. Follow-up? - forget it! The School Counselor, 1970, 17 (3), 228-232.

Krueck, T. Suggested procedure for conducting a student follow-up. Report produced by the Department of Research and Evaluation, Career Development Center, Skyline High School, Dallas, Undated.

Nelson, Jack L. Follow-up study of graduates. Improving College and University Teaching, 1964, 12 (2), 111-112.

Parton, M. Surveys, Polls, and Samples. New York: Harper & Row, 1950, Chap. 1.

Reuss, Carl. F. Differences between persons responding and not responding to a mailed questionnaire. American Sociological Review, 1943, 8, 433-438.

Rothney, John W. M. What are high school graduates doing ten years after high school? Vocational Guidance Quarterly, 1964, 13 (2), 134-136.

Rothney, John W. M. and Mooren, R. L. Sampling problems in follow-up research. Occupations, 1952, 30, 573-578.

Sluttleworth, Frank. Sampling errors involved in incomplete returns to mail questionnaires. Journal of Applied Psychology, 1941, 25, 588-591.

Snyder, Eldon W. A longitudinal analysis of the relationship between high school student values, social participation, and educational-occupational achievement. Sociology of Education, 1969, 42 (3), 261-270.

Texas Education Agency (TEA). Texas Education Product Study: Digest of Final Report. Department of Occupational Education and Technology, Texas Education Agency, Austin, Texas 1971.